Neuron, Vol. 21, 1235–1238, December, 1998, Copyright ©1998 by Cell Press

Minireview

Coincidence Detection in the Auditory System: 50 Years after Jeffress

Philip X. Joris,* Philip H. Smith,† and Tom C. T. Yin‡§

*Division of Neurophysiology
K.U.Leuven Medical School
Leuven, B-3000
Belgium

†Department of Anatomy
‡Department of Physiology
University of Wisconsin
Madison, Wisconsin 53706

The idea that cortical neurons code information not only in their average rate of discharge but also in the temporal pattern of spike discharge has generated much recent controversy (Shadlen and Newsome, 1998; Stevens and Zador, 1998). Proponents of time codes often use examples in organisms portrayed as being "specialized" for temporal processing (e.g., electric fish, barn owl, echolocating bats; see Carr, 1993). It is less widely known that even in "nonspecialized" mammals, including humans, a well-documented case of information encoding based on submillisecond spike timing is found in the auditory brainstem.

The role of the time dimension in hearing has been debated for over a century, and the extent to which perception of certain sound attributes (e.g., pitch; Cariani and Delgutte, 1996) depends on fine timing of firing patterns in the auditory nerve is still unsettled. However, a clear example of the importance of the temporal dimension comes from binaural hearing. The interaural time difference (ITD) between the arrival time of acoustic energy at the two ears provides an important cue for the spatial location of sound and can be discriminated at very small values ($\sim$10 μs). Physiologically, neurons in the medial superior olive (MSO) are sensitive to microsecond differences in their afferent signals. This is one of the few sensory circuits in the mammalian CNS for which a strong functional hypothesis can be formulated and the mechanisms underlying its physiological properties are relatively well understood.

In a seminal short paper published exactly 50 years ago, Jeffress (1948) outlined a model to transform acoustic timing differences into a neural spatial gradient, in effect transforming a time code into a space code. To a remarkable extent, the key assumptions of his model have been borne out, even though they were formulated at a time when single cell physiology of the brainstem was in its infancy. The model has guided much subsequent psychophysical and physiological research into the mechanisms of interaural time comparisons and is currently the backbone of virtually all computational models of sound localization (Colburn, 1996).

Jeffress made three assumptions (Figure 1A). First, monaural channels ("secondary fibers"), which originate from left and right cochlear nuclei, converge on a binaural nucleus ("tertiary fibers") and contain temporal information in their discharge pattern about the waveform of the acoustic stimulus. Second, cells in the binaural nucleus only discharge when receiving coincident spikes from their monaural, excitatory afferents. Third, the monaural channels project with ladder-like branching patterns running in opposite directions to the tertiary cells. In Figure 1A, the length of the ipsi- and contralateral inputs to the third order cells is equal up to points X and Y. Branch "a" forms a set of collaterals so that the axonal length to the tertiary cell is longer for cell 7 than for cell 1 and vice versa for the branch from the left side. Due to the finite axonal conduction speed, these branching patterns set up delay lines that run in opposite fashion for the two monaural input channels. We refer to the differences in delay between ipsi- and contralateral signals at their points of convergence as internal delays.

Imagine a sound originating from straight ahead. It reaches the cochleae without interaural delay (ITD = 0 μs). Only cell 4 receives coincident input signals and will be activated, since the total pathlengths traversed by its ipsi- and contralateral inputs are equal with an internal delay of 0 μs. If the tone originates from the left hemifield, it reaches the left ear earlier than the right. Input signals coincide if the signals from the left have to travel a longer path length than those from the right by an amount that exactly offsets the acoustic delay, e.g., at cell 1. Thus, the overall result of this scheme is the creation of a spatial array of cells, each tuned to a specific ITD and arranged orderly according to the azimuth to which they are tuned.

Modern anatomical and physiological evidence shows that the MSO is the likely site of Jeffress' model. It consists of a long sheet of cells oriented in a roughly parasagittal plane. Afferents from bushy cells in the anteroventral cochlear nucleus (AVCN) are segregated through the cells' bipolar dendritic morphology: ipsilateral afferents terminate on the lateral, contralateral afferents on the medial dendrites. What, then, is the current evidence in the MSO circuit for the three key assumptions of the Jeffress model and its main outcome, a spatial map of ITD?

Synchronization of Discharge Pattern to the Acoustic Stimulus

Clearly, to encode ITDs, it is necessary that the temporal information in the acoustic signal be encoded by the cochlea, as evidenced by phase locking to low-frequency tones in auditory nerve fibers (Figure 2A). Recent evidence (Joris et al., 1994) shows that phase locking in bushy cells is much more precise than in auditory nerve; at low frequencies ($<$1 kHz), spikes lock to a narrower range of phases and discharge on every stimulus cycle (Figure 2B). This higher phase locking reduces the temporal window over which binaural coincidences can occur and thereby sharpens ITD tuning in the MSO. It may derive from a monaural coincidence mechanism, reflecting the remarkable physiological and anatomical features of bushy cells (see Young, 1998).

§To whom correspondence should be addressed (e-mail: yin@physiology.wisc.edu).

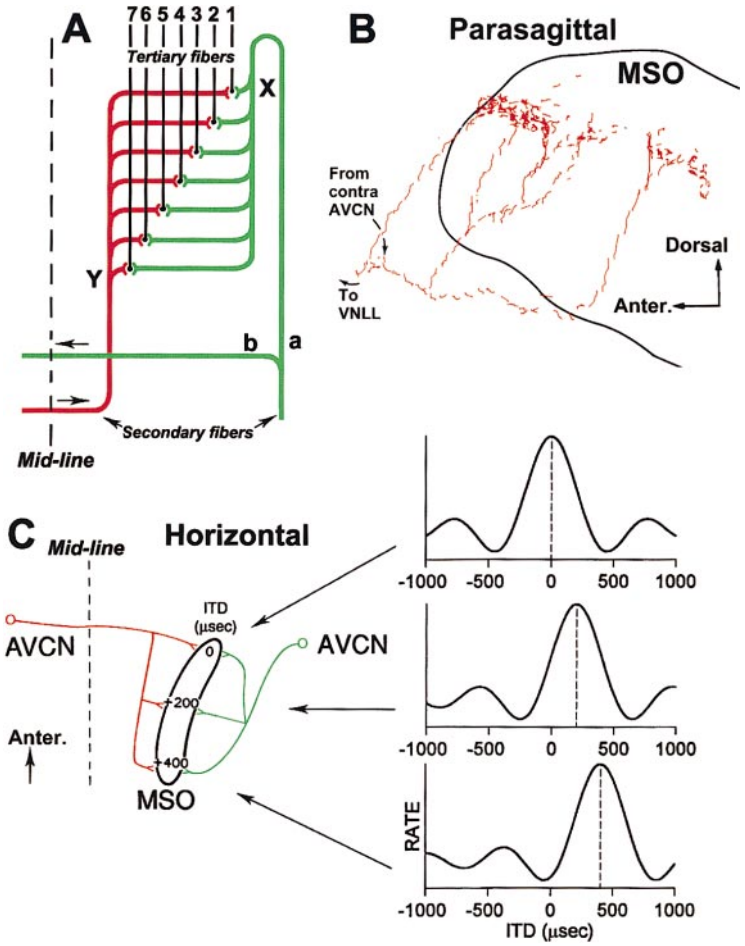

Figure 1. Anatomical Features of the Jeffress Model

(A) The original Jeffress (1948) delay line/coincidence model shown for the right side of the brainstem.

(B) Delay line configuration of a bushy cell axon (red) from the contralateral AVCN projecting to the MSO (black). Terminal branching occurs in a narrow iso-frequency strip. Adapted from Smith et al. (1993).

(C) Current view of the Jeffress model in the cat. Monaural channels feed into a binaural processor consisting of a bank of cross-correlators that each tap the signal at a different ITD. The cells for which the internal delay exactly offsets the acoustic ITD are maximally active.

Anatomical Evidence for Delay Lines

Intraaxonal recordings from individual bushy cell axons in cats with subsequent horseradish peroxidase (HRP) injection and histological reconstruction (Smith et al., 1993) showed that single afferents project to both ipsi- and contralateral MSO. Figure 1B is a computer reconstruction of the projection of one axon to the contralateral MSO. The axon approaches the MSO rostrally, turns caudally, and gives off branches in a fashion much like the hypothesized delay line of Jeffress. This arborization pattern, with more axonal length interposed between the parent axon and caudal branches than rostral branches, was found in most of the labelled contralateral MSO afferents. By contrast, in ipsilateral MSO afferents the pattern was much more varied, and there was no convincing delay line configuration or rostrocaudal bias.

These anatomical results are only partially consistent with the Jeffress model, which contains a dual delay line configuration (Figure 1A). Note, however, that the model still works with a single delay line. The results are also consistent with evidence from gross HRP injections in birds. The configuration of afferents to nucleus laminaris in chicks and barn owls (Young and Rubel, 1983; Carr and Konishi, 1990) is like that in the cat, with a ladder-like branching pattern in contralateral and an equally forked configuration in ipsilateral afferents. However, delay lines in the barn owl have been proposed on the basis of extracellular field potentials not to depend on this branching pattern but rather on conduction delays incurred in the terminal branches inside nucleus laminaris (Carr, 1993).

Coincidence Detection

Although straightforward in principle, there are few experimental tests of coincidence detection in the MSO (Goldberg and Brown, 1969) due to the difficulty in obtaining good recordings. The response of MSO cells is strongly modulated by ITD with a periodicity matching that of the stimulus (Yin and Chan, 1990). In response to monaural stimuli, the cell in Figure 2D phase locks to either ear, but the contralateral response is slightly delayed (by 140 μs) relative to the ipsilateral response. Thus, to binaural stimulation, we expect coincident afferent signals and a maximum response in this cell when the stimulus to the ipsilateral ear is delayed by ~140 μs, as seen indeed in the ITD function (Figure 2C). All MSO cells showed similar agreement between binaural and monaural responses (Yin and Chan, 1990). Little is known about the cellular mechanisms that allow MSO cells to perform such high-resolution coincidence detection, except that they have a low-threshold potassium channel (Smith, 1995) that limits temporal integration of bilaterally stimulated inputs (Grothe and Sanes, 1994).

Even though these data fit the predictions of the Jeffress model, some features of the responses are unexpected. Why would a coincidence detector discharge to monaural stimuli? Bushy cells have spontaneous activity; thus "driven" spikes from the stimulated ear can

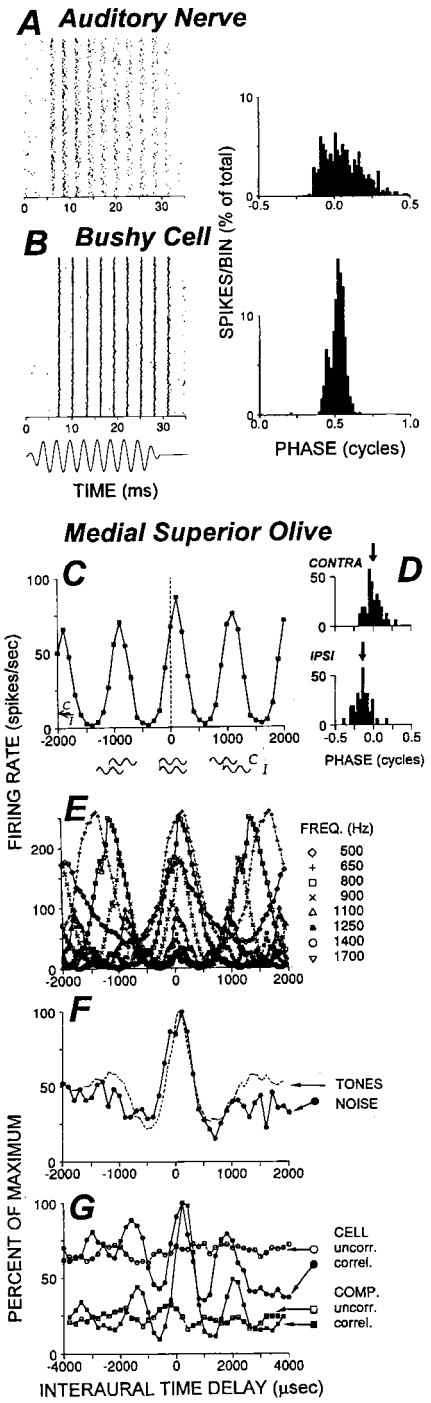

Figure 2. Physiological Features of the Jeffress Model

(A and B) Comparison of stimulus synchronization in an auditory nerve fiber and AVCN bushy cell. Each dot in the rasters (left panels) indicates a spike occurrence to a short tone at the cell's best frequency (350 Hz in [A] and 340 Hz in [B]); each row of dots is the response to one of the 200 repetitions. The responses are phase locked to the stimulus, as seen in the tendency of spikes to occur at a particular phase angle by graphing the response relative to stimulus phase (right panels). Spikes in the bushy cell are temporally less dispersed than in the auditory nerve and occur in each stimulus cycle, whereas cycles are often skipped in the nerve. Adapted from Joris et al. (1994).

(C–G) Sensitivity to ITD in MSO neurons.

(C) Responses to a binaural 1 kHz tone as a function of ITD. Positive ITDs represent a lead of the contralateral (C) relative to the ipsilateral (I) tone.

(D) Period histograms of the same cell to the same monaural ipsi- or contralateral 1 kHz tone show average response phases (arrows) that differ by 0.14 cycles, or 140 μs.

(E and F) Delay functions for another MSO cell to tones at many different frequencies (E) and to a broadband noise stimulus (F). Dashed line in (F) is the linear addition of responses in (E).

(G) Responses of a third MSO cell were obtained to correlated (closed circles) and uncorrelated (open circles) noise over a range of ITDs, as well as to these same stimuli presented monaurally. The monaural responses were then cross-correlated at different time delays, for the correlated (closed squares) and uncorrelated (open squares) waveforms. Adapted from Yin and Chan (1990).

coincide with "spontaneous" spikes from the unstimulated ear to generate a low monaural response. Also, there can be monaural coincidences by virtue of the convergence of multiple homolateral afferents on the MSO cell. Agmon-Snir et al. (1998) proposed a biophysical model in which the segregation of ipsi- and contralateral afferents on different dendrites improves ITD detection by boosting the effect of binaural over monaural coincidences.

Because sinusoids are perfectly periodic, the input signals are coincident at multiple ITDs and delay functions are periodic (Figure 2C). However, there is only one ITD, the so-called characteristic delay (Rose et al., 1966; Yin and Kuwada, 1983), where the same degree of coincidence is reached independent of the frequency content of the acoustic stimulus, and the acoustic delay in the stimulus exactly opposes the internal delay. Figure 2E shows delay functions at many different frequencies. This cell consistently shows a maximum in its response at only one ITD (~30 μs). The response to a wide-band noise stimulus (Figure 2F, closed circles) shows only one peak at this same ITD at which all frequency components push the response in the same direction, as predicted from linear summation of the tonal responses (Figure 2F, dashed line).

Besides being wideband, the spectrum of natural sounds is often also different for the two ears. How can the concept of coincidence detection be extended to make quantitative predictions of responses to such stimuli? Basically, each coincidence detector is sensitive to similarity in the time structure of ipsi- and contralateral spike trains at a particular time delay. This is akin to the mathematical operation of cross-correlation, which is used in most computational models to implement coincidence detection of the input signals after bandpass filtering and rectification by the cochlea (Colburn, 1996). The ability to independently control the input signal to the two ears, and thereby the firing patterns in the afferents to the MSO, enabled thorough testing of the cross-correlation formulation (Yin et al., 1987). For example, if an identical noise signal is presented to the two ears at various ITDs, the response of the cell is clearly dependent on ITD (Figure 2G, closed circles, correl.) but not when two uncorrelated noise tokens are used (open circles, uncorrel.). Independent confirmation that the cell computes a cross-correlation is again based on analysis of the monaural responses. Spike trains to ipsi- and contralateral monaural presentations of the noise stimuli were obtained and were computationally

cross-correlated. The computed output was independent of ITD for monaural responses to uncorrelated noise (open squares, uncorrel.) and dependent on ITD for responses to correlated noise (closed squares, correl.), with features much like the actual binaural responses.

A Map of ITDs

The result of Jeffress' model is a representation of ITD as a spatial activity profile in an array of tuned cells. Unfortunately, this part of the model has been the most elusive to document. Best ITDs in the MSO are strongly biased toward the contralateral hemifield and fall mostly within the range of ITDs naturally encountered by a cat (0–400 μs) (Yin and Chan, 1990). Moreover, there is a significant correlation between the best ITDs and the rostrocaudal location of recording sites pooled from different experiments. The most rostrally located cells prefer small ITDs, while cells located caudally prefer large positive ITDs as generated by sounds in the contralateral hemifield (Figure 1C). However, the relationship is quite scattered and needs to be reexamined by methods that do not necessitate pooling across animals. Importantly, it is in agreement with the anatomical findings on contralateral delay lines extending rostral to caudal, which predict best ITDs to be small rostrally and large caudally. This observation departs from the original scheme of Jeffress, in which ITDs corresponding to both ipsi- and contralaterally located sources are encoded on both sides of the brain, but it is consonant with the contralateralization generally found in other sensory and motor systems.

In conclusion, the current view on binaural time processing contains the essential features proposed by Jeffress. Just as Hubel and Wiesel's model of orientation selectivity has provided a bull's eye for vision research, the simple model introduced by Jeffress has served as a focal point in audition and has sparked many computational and experimental studies, illustrating the heuristic value of such qualitative models in neuroscience. To end, we mention one particularly interesting study that bridges the gap between human psychophysics and experimental studies in animals. Multiple sclerosis patients with focal lesions in the pons showed abnormal discrimination of ITDs, as well as abnormalities in their brainstem evoked responses, but could still discriminate interaural intensity differences (Levine et al., 1993). These findings were interpretable in the framework of Jeffress' model by assuming that demyelination caused a desynchronization or decorrelation of the spike patterns fed to the binaural processor. Such decorrelation impairs discrimination of ITDs but not of intensity differences, which are usually thought to require a comparison of average spike counts and to be unaffected by disruption of fine time structure. While this study does not bear directly on the details of the Jeffress model, it provides a compelling argument for the importance of spike timing in human auditory functioning.

Selected Reading

Agmon-Snir, H., Carr, C.E., and Rinzel, J. (1998). Nature *393*, 268–272.

Cariani, P.A., and Delgutte, B. (1996). J. Neurophysiol. *76*, 1698–1716.

Carr, C.E. (1993). Annu. Rev. Neurosci. *16*, 223–243.

Carr, C.E., and Konishi, M. (1990). J. Neurosci. *10*, 3227–3246.

Colburn, H.S. (1996). In Auditory Computation, H.L. Hawkins, T.A. McMullen, A.N. Popper, and R.R. Fay, eds. (New York: Springer), pp. 332–400.

Goldberg, J.M., and Brown, P.B. (1969). J. Neurophysiol. *32*, 613–636.

Grothe, B., and Sanes, D.H. (1994). J. Neurosci. *14*, 1701–1709.

Jeffress, L.A. (1948). J. Comp. Physiol. Psychol. *41*, 35–39.

Joris, P.X., Carney, L.H., Smith, P.H., and Yin, T.C.T. (1994). J. Neurophysiol. *71*, 1022–1036.

Levine, R.A., Gardner, J.C., Stufflebeam, S.M., Fullerton, B.C., Carlisle, E.W., Furst, M., Rosen, B.R., Kiang, N.Y.S. (1993). Hear. Res. *68*, 59–72.

Rose, J.E., Gross, N.B., Geisler, C.D., and Hind, J.E. (1966). J. Neurophysiol. *29*, 288–314.

Shadlen, M.N., and Newsome, W.T. (1998). J. Neurosci. *18*, 3870–3896.

Smith, P.H. (1995). J. Neurophysiol. *73*, 1653–1667.

Smith, P.H., Joris, P.X., and Yin, T.C.T. (1993). J. Comp. Neurol. *331*, 245–260.

Stevens, C.F., and Zador, A.M. (1998). Nat. Neurosci. *1*, 210–217.

Yin, T.C.T., and Chan, J.C.K. (1990). J. Neurophysiol. *64*, 465–488.

Yin, T.C.T., and Kuwada, S. (1983). J. Neurophysiol. *50*, 1020–1042.

Yin, T.C.T., Chan, J.C.K., and Carney, L.H. (1987). J. Neurophysiol. *58*, 562–583.

Young, E.D. (1998). In The Synaptic Organization of the Brain, G.M. Shepherd, ed. (Oxford: Oxford University Press), pp. 121–158.

Young, S.R., and Rubel, W.E. (1983). J. Neurosci. *3*, 1373–1378.